SOCIAL SECURITY'S EARNINGS TEST PENALTY AND THE EMPLOYMENT RATES OF ELDERLY MEN AGED 65 TO 69

Stephen Rubb
Bentley College

INTRODUCTION

Social Security provides retirement income to eligible elderly individuals who reach age 62 and apply for benefits. Beyond this age, some recipients continue to work on a full- or part-time basis. The Social Security Administration reduces the annual level of benefits for those recipients who have earnings above a specified amount known as the earnings test threshold. In 1990 the earnings test threshold was $9,360 for individuals aged 65 to 69. The threshold is adjusted annually to account for wage inflation. In 1989, Social Security's earnings test reduced retirement benefits for 926,000 recipients by over $5 billion and removed all benefits for some 333,000 individuals [Bondar, 1993]. Effective in 1990, the earnings test penalty for a person aged 65 to 69 was reduced from 50 cents to 33 cents for every dollar earned in excess of the annually adjusted threshold.

According to Friedberg [2000, 48] the labor supply response to the 1990 reduction in the earnings test penalty is an area that is "mostly unexplored." Additionally, the "literature does not address one question of substantial interest, which is whether the elderly will be more likely to work at all as the earnings test is loosened" [Gruber and Orszag, 2000, 10]. The research presented in this paper attempts to fill this void. This is the first known empirical study to focus solely on the impact of the 1990 rule change on the employment decisions of the elderly.[1] By focusing on the 1990 change in law, this paper is able to estimate non-parametric models that compare outcomes of a treatment group with various control groups. With limited exception, the literature (discussed below) concludes that the impact of the earnings test on annual hours of labor supply conditional on employment is small. Yet clear evidence of the impact of the earnings test on behavior can be found among wage earners clustered near the threshold [Leonesio, 1990; 1993; Bondar, 1993; Friedberg, 2000]. It is important for public policymakers to understand the impact of these rule changes on the employment decisions of the elderly because of their tremendous impact on governmental budgetary concerns and on the behavior and well being of the elderly.

This paper begins by describing some of the theoretical expectations of employment and labor supply responses to the 1990 reduction in the earnings test penalty. This is followed by a review of the previous literature and an examination of the

Stephen Rubb: Department of Economics, Bentley College, Waltham, MA 02452-4705. E-mail: srubb@bentley.edu.

long-term decline in labor force participation rates of elderly men aged 65 to 69. A graphical analysis and an unadjusted difference-in-differences table with various control groups are also presented. Finally, empirical models are used to examine the impact of the rule change on the employment decision and annual hours worked.

THEORETICAL EXPECTATIONS

The theoretical expectations of a reduction in the earnings test penalty are multifaceted. Two areas are worth examining. First, the earnings test may impact the number of hours of labor supplied, conditional on employment. The theoretical impact on hours worked is complex and ambiguous due to offsetting substitution and income effects. Second, economic theory suggests that lowering the earnings test penalty should lead to an increase in employment rates among the elderly through a substitution effect. The substitution effect will be particularly responsive if a significant hours-constraint on workers exists.

The earnings test results in a kinked budget constraint. A lower earnings test penalty results in a net marginal wage rate[2] increase only for those with earnings above the threshold who collect Social Security benefits. Among those who work, two groups may not be directly affected by the rule change at all. The first group is individuals whose "likely earnings" (earnings that would have occurred with the previous 50 percent earnings test penalty) are below the threshold. Over 65 percent of working beneficiaries during the survey years (1984-1993) do not have benefits that are directly impacted by the earnings test because their earnings are below the threshold.[3] The labor supply decision will be indirectly impacted if employees choose a wage-hours package so that earnings are kept below the threshold. The second group is individuals whose likely earnings exceed the threshold by three times the amount of potential Social Security benefits; the earnings test penalty of 33 percent (or 50 percent) make them ineligible for benefits. Individuals with likely earnings either below or substantially above the threshold experience no change in Social Security benefits or net marginal wage rates.

Between these two extremes lies the complexity of determining the theoretical impact of the 1990 rule change on annual hours of work conditional on employment. Individuals with likely earnings at (or just above) the threshold may be directly influenced by the reduction in the earnings test penalty. At this level of earnings the income effect will not be realized. Net marginal wage rates, however, increase by 33 percent. The substitution effect predicts that in the absence of an hours constraint these individuals will work additional hours. The opposite occurs for individuals whose likely earnings exceed the threshold by two to three times the amount of their potential Social Security benefits. Such individuals would receive no benefits with a 50 percent earnings test penalty, but are eligible to receive some (albeit reduced) benefits with a 33 percent earnings test penalty. Ironically, now that the individuals collects Social Security benefits, net marginal wage rates are reduced. Both the income and substitution effects predict a decline in annual hours of labor supplied. In summary, the theoretical expectation of the impact of a reduction in the earnings test penalty on hours worked among those with earnings above the threshold is ambiguous.

In analyzing the employment decision of 65- to 69-year olds, it is important to note the following. No income effect exists if the beneficiary chooses not to work. The substitution effect indicates that higher net average wage rates should encourage employment particularly if workers have hours constrained by a minimum hours requirement or there are an insufficient number of acceptable part time jobs.[4] Reducing the earnings test penalty may induce people to accept such employment that would presumably push earnings above the threshold. Given these constraints, labor supply theory predicts an unambiguous increase in the number of employed workers, *ceteris paribus*. However, the impact of the 1990 rule change on net average wage rates may be small because it only impacts earnings in excess of the threshold. Furthermore, the marginal wage rate for the first hour of labor supplied as an individual enters the labor market remains unchanged.

PREVIOUS LITERATURE AND EMPLOYMENT TRENDS

Blinder, Gordon, and Wise [1980; 1981] note that many of Social Security's other provisions, such as the delayed retirement credit (see Appendix) and benefit adjustment through increases in average lifetime earnings (and contributions) combined with the earnings test, may actually encourage work. In a similar vein, Gustman and Steinmeier [1991] run labor simulations involving the earnings test penalty and other Social Security provisions and find the earnings test to be unimportant for 65- to 69-year olds. In some respects, these findings mask the impact of the earnings test with other Social Security provisions.

The number of older males with earnings above the threshold is relatively small and declining. Some authors [Burtless and Moffitt, 1984; Burtless and Moffitt, 1985; Packard, 1990; Honig and Reimers, 1989] suggest that this low number indicates that changes in the earnings test penalty will only impact a few people; thus the aggregate labor supply effect of changes in the earnings test will be small. Packard predicts that "at least 80 percent, and perhaps more than 90 percent, of ... persons age 65 to 69 will not change their level of work effort if the earnings test is eliminated (as was done in 2000 for individuals age 65 to 69)" [1990, 2]. A counter interpretation of the same low number suggests that the reason the number is low is that those affected could be endogenous to the mere existence of the earnings test penalty.

Reimers and Honig [1993; 1996] examine how retirees respond to changes in future benefits in a model that examines workforce re-entry in the 1970s. Their primary finding is that men respond to current benefits rather than adjustments to future benefits (see discussion on the delayed retirement credit in the Appendix). With respect to the research of this paper, they find that a 10-percent increase in the earnings test threshold increases the likelihood of re-entry into the work force by 20 percent. They conclude "that the budget constraint is discontinuous, due to either fixed costs of participation or scarcity of acceptable part-time jobs" [Reimers and Honig, 1993, 201].

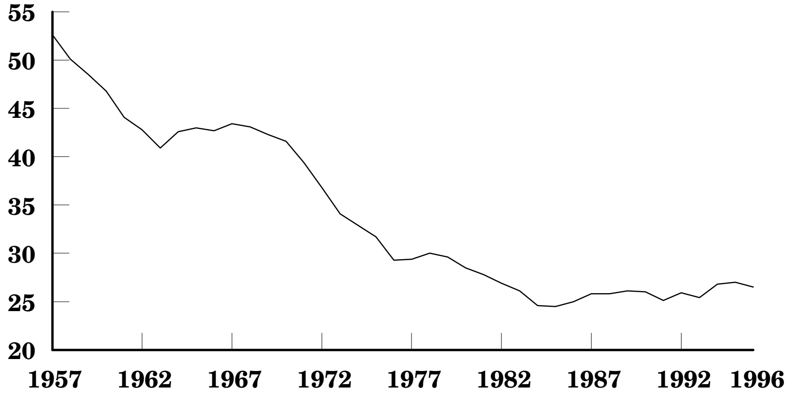
Gruber and Orszag claim that the earnings test exerts "no robust influence on the labor supply decisions of men" [2000]. To partially support this claim, their analysis includes a graphical comparison between individuals aged 66 to 69 and those aged 63 to 64. Their graphs indeed fail to capture any noticeable difference in the employment rates between the two groups after the 1990 rule change. Their econometric models capture changes in the earnings test threshold, exemption age, and penalty rate that occurred over a twenty-five year period for a broad range of age/year groupings (not just 65- to 69-year olds) on a yearly basis at the macro level. One of their dependent variables is the employment rate within each of the 338 age/year groupings.

With regard to research presented in this paper, Gruber and Orszag's independent variable of interest is "tax rate change" (earnings test penalty change). Two earnings test tax rates exist. Groups of 66- to 69-year olds in the 1990s have a 33 percent tax rate. Groups of 63-to 64-year olds, individuals aged 71 prior to 1983, and elderly aged 66 to 69 prior to 1990 have a 50 percent tax rate.[5] Also included are several groups exempt from the earnings test or not eligible for Social Security. This includes groups of 59- to 61-year olds (not eligible for Social Security), 73- to 75-year olds (exempt), and 71-year olds since 1983 (exempt). Using this larger context, they conclude that the impact of changes in the earnings test penalty on employment, if any exists, will be small.

Gruber and Orszag find changes in the earnings test (threshold, exemption age, and penalty rate) to have a relatively strong impact on a dependent variable measuring the percentage of individuals in each group who collect benefits. The theoretical expectations are straightforward. The existence of the earnings test results in some beneficiaries with high earnings becoming ineligible to collect benefits, thus a negative relationship is predicted (and found). A higher threshold increases the number of working beneficiaries with earnings low enough to be eligible to collect benefits, thus a positive relationship is predicted (and found). A higher earnings test penalty lowers the number of individuals eligible to collect benefits. For example, an individual with earnings that exceed the threshold by 2.5 times his (potential) Social Security benefits will be able to collect some benefits under a 33 percent penalty, but none under a 50 percent penalty. Gruber and Orszag find changes in the penalty to have a negative but statistically insignificant impact on the rate of individuals claiming benefits.

Friedberg notes "the working elderly bunch in substantial numbers at and just below" the threshold and strongly argues that this "cast(s) doubts on earlier findings that the earnings test has little impact on behavior" [2000, 48]. One perplexing finding is that the amount of bunching near the threshold after the 1990 rule change (which she examines graphically) does not fall. Friedberg's econometric analysis examines the impact of the 25 percent increase in the threshold for individuals aged 65 to 71 in 1978 and of the removal of the earnings test penalty for individuals aged 70 and 71 in 1983 on aggregate hours. Individuals ages 62, 65, 70, and 72 are not included in the study for technical reasons. Using estimates of the income and substitution elasticities, Friedburg calculates that eliminating the earnings test will increase aggregate hours by 5.3 percent among individuals aged 65 to 69. The framework used by Friedburg does not allow for an examination of the employment decision.

FIGURE 1
The Annual Average Civilian Labor Force Participation Rates Of Elderly Men Aged 65 to 69 in the U.S. 1957 to 1996



Source: U.S. Bureau of Labor Statistics, Employment and Earnings.

Figure 1 shows the annual average civilian labor force participation rates of men aged 65 to 69 in the United States from 1957 to 1996.[6] Rates fell from over 50 percent in 1958 to below 25 percent in 1984. In fact, rates had been falling for 50 years prior to the introduction of Social Security in the 1930s. Despite the observed decline, the impact of the earnings test penalty on labor force participation is uncertain. The decline could be due to increases in the real value of Social Security benefits and other retirement incomes, not the earnings test penalty. The data also suggests that the labor force participation rates have been edging slightly upward since the late 1980s [Herz, 1995].

An ordinary least squares regression is used to determine the impact of time trends, unemployment rates, and a time shift variable on participation rates. The functional form of the model is the following:

$$LFPR_t = \alpha + \beta(TIME_t) + \gamma(UR_t) + \rho(D_t) + \varepsilon$$

LFPR is the labor force participation rate in year t. *TIME* is a time trend variable (where *TIME* = 1 in 1957). *UR* is a variable representing the state of the economy as indicated by the national unemployment rate. *D* is a dummy variable that takes on the value of one for the years 1990 and beyond. The expected sign of β is negative due to the long-term trend of declining labor force participation rates of elderly men. The expected sign of γ is negative because higher unemployment rates generally lead to lower participation rates through a discouragement effect. The sign of ρ should be positive if there is a reversal in the downward trend since 1990.

The results of the regression are as follows:

$$LFPR_t = 55.2 - \underset{\mathit{20.7}}{0.77}(TIME_t) - \underset{\mathit{4.7}}{1.07}(UR_t) + \underset{\mathit{6.0}}{6.20}(D_t)$$

The *t*-statistics are in italics. The adjusted R-squared is .96. The F-statistic is 290.4 and significant at the .01 level. The findings suggest that a higher unemployment rate does moderately depress the labor force participation rates of older men and that the decline has reversed significantly since 1990. A possible explanation of the reversal could be the 1990 change in the earnings test penalty. However, the increase in participation rates since 1990 can possibly be explained by other variables such as improved labor market conditions following the end of the recession in early 1991, changes in the educational attainment levels of the elderly, and changes in private pensions.

EMPIRICAL STRATEGY

The data used in this study came from the annual demographic supplements of the March 1985 to the March 1994 *Current Population Surveys* (CPS). The CPS public use data set contains considerable micro level data. I chose 1984 to 1993 to avoid implications of changing policies regarding the taxation of Social Security benefits. Prior to 1984, Social Security benefits were not taxed. Between 1984 and 1993 up to 50 percent of one's Social Security benefits were considered taxable income. Since then, up to 80 percent of an individual's Social Security benefits are considered taxable income. CPS data reflects an individual's age during March of the following year. As is noted in Gruber and Orszag [2000], this leads to ambiguity about the individual's actual age during the year in question. Accordingly, I take a conservative approach and include only individuals aged 66 to 69.

Women and minority males are excluded. As pointed out by Bondar [1993], women accounted for less than one-third of 65- to 69-year olds directly affected by the earnings test. Additionally, women of this generation observed significantly lower average hourly wage rates than men. As a result, fewer women are impacted by the mere presence of the earnings test. Likewise, minority males are excluded because the lower wages of this cohort tends to complicate the analysis. According to wage estimates from the survey, the average minority male could work 50 percent more hours per year than the average non-Hispanic white male before the earnings test would come into play. A small number of individuals who received a federal or state government pension and did not receive any Social Security benefits are believed to be ineligible for Social Security and excluded from the analysis. Lastly, the sample excludes individuals with negative personal and family non-labor incomes in order to compute the natural logs of these variables plus one. After these exclusions, the sample consists of 17,572 non-Hispanic, white males, aged 66 to 69, who were likely eligible for Social Security benefits (34,653 males aged 63 to 72).

Probit models of the employment behavior of 66- to 69-year-old white men are estimated with a dependent variable of one or zero based on whether an individual worked or not during a given year. The models take the functional form:

TABLE 1
Description of Variables

Variables	Mean Values for Age 66-69	63-64	71-72	Description of Variables
AGE	5.45	1.50	9.48	own age minus 62
ED8	0.19	0.16	0.22	1 if educational attainment through grade 8
ED11	0.15	0.15	0.16	1 if educational attainment grade 9 through 11
ED13	0.14	0.15	0.14	1 if educational attainment includes some college
ED16	0.09	0.11	0.09	1 if educational attainment is 4 years of college
ED17	0.08	0.09	0.07	1 if educational attainment is 5+ years of college
MARRIED	0.82	0.83	0.81	1 if married with spouse present
METRO	0.49	0.50	0.49	1 if residence in a SMSA
LN~FY	7.32	7.10	7.27	other family income; excludes own earnings & SS benefits
LN~PY	7.24	6.80	7.29	other personal income, excludes own earnings & benefits
STATEUER	6.34	6.44	6.38	annual state unemployment rate
WAGE	2.02	2.17	1.94	imputed of hourly wage rate (natural log)
YEAR	5.41	5.34	5.55	time trend variable (1984=1)
A66~9	1.00	0.00	0.00	1 if age 66 to 69
Y90~3*A66~9	0.39	0.00	0.00	1 if year is 1990 to 1993 and aged 66 to 69
Y90~3	0.39	0.38	0.41	1 if year is 1990 to 1993

Wage and incomes in 1983 dollars. Natural log of income plus one is used. Wage estimate based on model including 63 to 72 year olds. Individuals with negative incomes excluded in all models.

$$Pr(W)_{it} = \alpha + \beta \mathbf{X}_{it} + \gamma Y90{\sim}93_{\mathrm{it}} + \rho YEAR_{it} + \varepsilon.$$

Pr(W) reflects the probability of work. **X** is a vector of independent variables described in Table 1 and discussed below. A time trend variable, *YEAR*, is included to capture the potential downward trend in employment rates. The dummy variable *Y90~93* equals one if the year is between 1990 and 1993. The impact of the 1990 rule change is expected to be captured with this variable. These single difference models test for a shift in levels of employment while maintaining the same secular time trend.

The independent variables include one's expected wage rate (*WAGE*, see Appendix) with an expected positive sign since a higher market wage should increase the likelihood that the market wage will exceed the reservation wage. Higher non-wage income (*LN~PY, LN~FY*) should result in a higher demand for leisure if leisure is a normal good. *LN~FY* will be positive if income of a spouse is complementary to personal income. Wage and income variables are expressed in logarithmic form to improve the fit of the models and to avoid the impact of outliers. Educational attainment is believed to be a negative predictor of employment when wages are included and a useful substitute for wages when wages are not included. As such all models include five educational variables (*ED8, ED11, E13, ED16, ED17*), with each representing a different meaningful level of schooling. Annual state unemployment rates (*STATEUER*) should be negatively related to the probability of working through a discouragement effect. Marital status (*MARRIED*, being married) could be negatively related to the probability of working if the presence of a spouse indicates a greater demand for leisure. Employment patterns tend to vary by geographic region. Thus the models include five regional dummy variables (not shown) and a variable indicating if an individual lives in a metropolitan area (*METRO*).

The possibility remains that unexplored factors outside the models are the primary cause of any observed change in the employment rates of elderly men aged 66 to 69. This includes, but is not limited to, improved health of the elderly, increased wealth of the general population, and changes in private pensions. To explore this hypothesis, difference-in-differences models are examined. That is to say, individuals aged 66 to 69 in the test group are pooled with individuals slightly younger (aged 63 and 64) and slightly older (aged 71 and 72) in various control groups. The models take the form:

$$Pr(W)_{it} = \alpha + \beta \mathbf{X}_{it} + \gamma Y90\text{~}93_{it} + \eta A66\text{~}69_{it} + \delta Y90\text{~}93_{it} {*} A66\text{~}69_{it} + \rho YEAR_{it} + \varepsilon$$

A66~69 indicates the individual is aged 66 to 69 and *Y90~93*A66~69* indicates an individual aged 66 to 69 from 1990 to 1993. If factors outside the models cause the increase in employment rates of those in the test group, it is likely that they cause a similar increase in the employment rates of the control group, both of which would impact γ. On the other hand, δ is believed to capture any unique impact the rule change has on the test group. An F test (the Chow Test) does not reject the hypothesis of no structural differences prior to 1990 between the two groups at the .01 level for both this model and the Tobit model (described next) when an imputed wage is included and the combined control group is used (63-, 64-, 71- and 72-year-olds).[7] The Chow test rejects the hypothesis when no wage term is included or when smaller control groups are used.

Tobit models capture the impact of the rule change on annual hours. The Tobit models have the identical specification as the Probit models but capture annual hours worked. By design, aggregate hours are not conditional on work because excluding individuals who do not work would induce sample selection bias in the estimates as is noted in Gruber and Orszag [2000]. The Tobit models are believed to be an appropriate specification of the annual hours estimates due to the censored nature of the data (roughly two thirds of the sample does not work).

All models are run twice, once with and once without an imputed wage variable to test for the strength of the wage estimates. Since the expected wage rate is a strong predictor of employment and aggregate hours worked, models that include the imputed wage variable are believed to be the appropriate specification. Lastly, the control group of individuals aged 63, 64, 71 and 72 is subdivided to check for the robustness of the findings. In total three control groups are used. Results for all models are given, but the discussion that follows focuses mainly on the models that include the imputed wage term and individuals aged 63, 64, 71 and 72 as the control group.

EMPIRICAL RESULTS

The data set is first used to estimate employment rates of individuals aged 66 to 69, 63 to 64, 71 to 72, and 63, 64, 71 and 72 included in the sample. Table 2 shows the

TABLE 2
Average Employment Rates for Select Age Groups (in sample)

Age	1984 to 1989	1990 to 1993	Difference in Differences
66 to 69	.336	.342	.007
Observations	10,774	6,798	17,572
Standard Error	(.472)	(.474)	(.473)
63, 64, 71, 72[a]	.418	.409	−.009
Observations	10,417	6,664	17,081
Standard Error	(.491)	(.487)	(.480)
Difference (with 66-69)	−.083	−.067	.016
63 to 64	.549	.533	2.016
Observations	5,954	3,600	9,554
Standard Error	(.498)	(.499)	(.498)
Difference (with 66-69)	−.213	2.191	.022
71 to 72	.214	.215	.001
Observations	4,463	3,064	7,527
Standard Error	(.410)	(.411)	(.410)
Difference (with 66-69)	.122	.127	.006

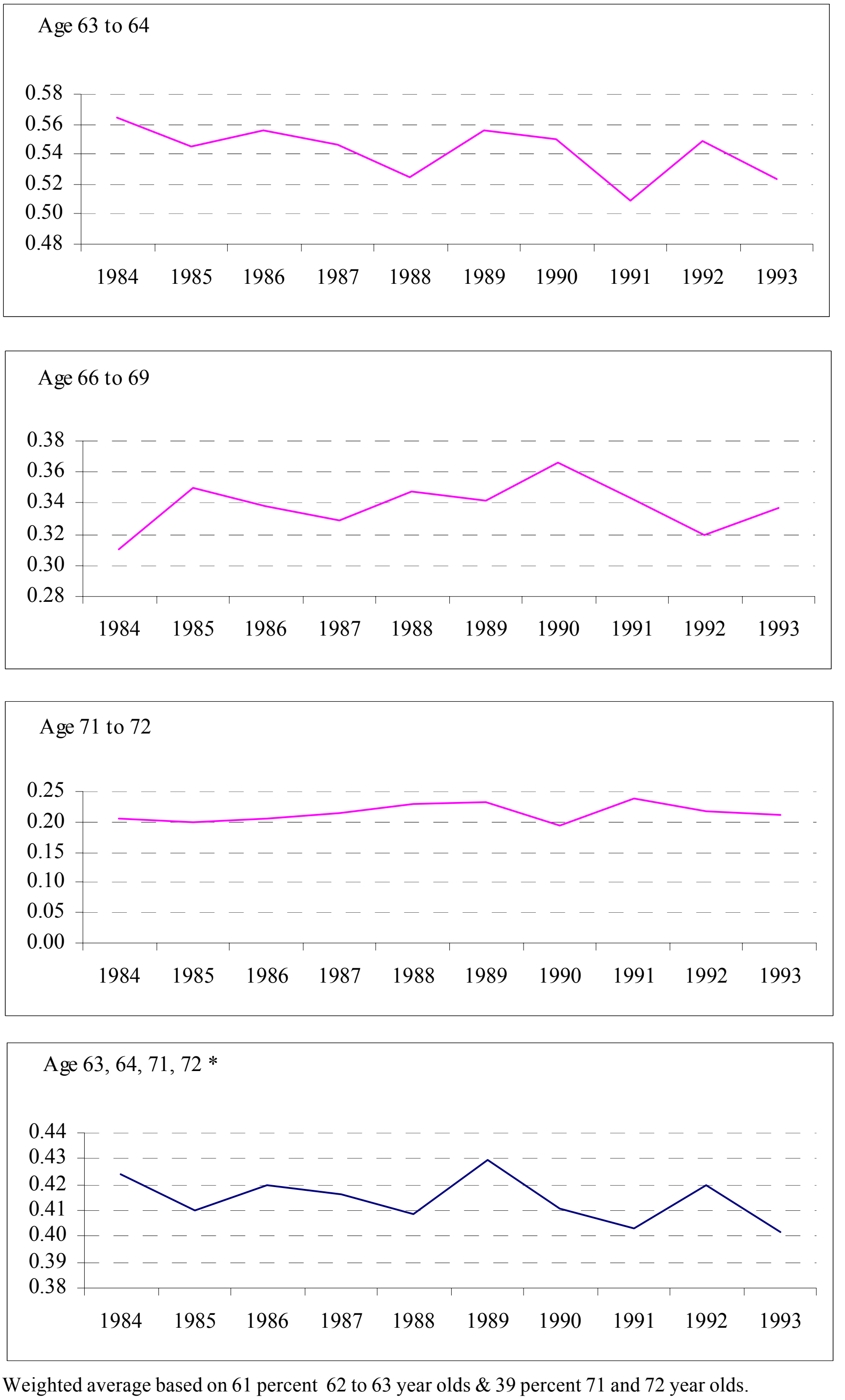
a. Weighted average based on 61- percent 62- to 63-year-olds and 39 percent 71 and 72 year olds, as observed in total sample.

average employment rates for the various age groups from 1984 to 1989 and from 1990 to 1993. The 0.7 percentage point increase in the employment rate of 66 to 69 year olds is consistent with the recent rise in labor force participation described above. Such a finding is also consistent with theoretical expectations based on the substitution effect. This compares to a 0.9 percentage-point decline in the combined employment rate of 63-, 64-, 71-, and 72-year-olds, a 1.6 percentage-point difference in the unadjusted differences. In a similar vein, Figure 2 and Figure 3 graph the annual average employment rates for each age group from 1984 to 1993. The graphs do not show any clear trend.

Table 3 and Table 4 present the findings of the eight Probit models and eight Tobit models, respectively. All models are found to be statistically significant at the .01 level. Most coefficients, including five regional dummy variables not shown, are found to be statistically significant at the .01 level. Each table shows the marginal effects of the respective variables. The marginal effects are calculated by multiplying the coefficients by the mean probability of employment in the sample.

The marginal effect for the time shift dummy, *Y90~93*, is .030 and is statistically significant at the .05 level when excluding a control group. This suggests a 3.0 percentage-point increase in the employment rate of elderly white men since 1990. This finding is consistent with the substitution effect, but may not reflect an increase in the probability of employment unique to the test group. The positive coefficients for

FIGURE 2
Annual Average Employment Rates for Select Age Groups by Year (in sample)



* Weighted average based on 61 percent 62 to 63 year olds & 39 percent 71 and 72 year olds.

TABLE 3
Probit Analysis of Probability of Employment

Variables	Age 66 to 69 & No Control Group		Control Group Age 63, 64, 71 & 72		Control Group Age 63 & 64		Control Group Age 71 & 72	
	No wage	Wage	No wage	Wage	No wage	Wage	No wage	Wage
C	.270	−2.157	.369	−2.851	.428	−3.289	.231	−1.854
	(9.9)	(28.2)	(20.8)	(48.8)	(19.3)	(48.0)	(7.6)	(27.0)
AGE	−.032	.006	−.039	−.008	−.042	.007	−.028	−.001
	(10.6)	(1.7)	(42.1)	(7.6)	(11.9)	(1.9)	(10.9)	(0.3)
ED8	−.092	.170	−.118	.238	−.136	.273	−.082	.139
	(9.0)	(13.2)	(14.7)	(23.3)	(13.6)	(21.9)	(10.6)	(13.7)
ED11	−.061	.068	−.081	.107	−.089	.114	−.058	.062
	(5.8)	(6.0)	(9.7)	(11.7)	(8.7)	(10.2)	(7.2)	(7.0)
ED13	.060	−.099	.061	−.170	.071	−.186	.050	−.094
	(5.7)	(8.4)	(7.4)	(18.0)	(7.0)	(16.0)	(6.2)	(10.2)
ED16	.161	−.279	.161	−.465	.193	−.529	.129	−.246
	(13.2)	(15.5)	(17.0)	(31.7)	(16.5)	(30.0)	(13.9)	(16.9)
ED17	.214	−.409	.240	−.653	.267	−.731	.193	−.366
	(16.3)	(18.0)	(23.3)	(34.8)	(20.9)	(33.4)	(19.2)	(18.9)
MARRIED	.038	−.123	.042	−.204	.055	−.213	.027	−.123
	(2.9)	(8.7)	(4.3)	(18.4)	(4.6)	(15.9)	(2.6)	(11.0)
METRO	−.032	−.211	−.027	−.274	−.025	−.319	−.030	−.174
	(4.5)	(23.4)	(4.8)	(38.0)	(3.7)	(35.9)	(5.6)	(25.0)
LN~FY	.004	.006	.005	.006	.005	.007	.004	.005
	(3.0)	(3.7)	(4.6)	(5.6)	(4.1)	(4.8)	(3.6)	(4.6)
LN~PY	−.024	−.006	−.028	−.008	−.034	−.008	−.019	−.006
	(20.8)	(4.8)	(31.6)	(7.8)	(31.1)	(6.5)	(21.9)	(6.7)
STATEUER	−.017	−.018	−.019	−.019	−.021	−.021	−.015	−.016
	(8.0)	(8.1)	(11.5)	(11.4)	(10.3)	(10.1)	(9.5)	(9.7)
WAGE	—	1.119	—	1.554	—	1.766	—	.963
		(33.7)		(57.5)		(57.1)		(33.7)
YEAR	−.007	.001	−.008	−.002	−.011	−.002	−.004	.000
	(2.9)	(0.2)	(4.3)	(1.0)	(4.7)	(0.9)	(2.5)	(0.0)
A66~69	—	—	−.029	.002	−.041	−.052	−.005	.052
			(4.2)	(0.2)	(2.5)	(3.1)	(0.4)	(3.9)
Y90~93	—	—	.005	.005	.007	.003	.002	.005
*A66~69			(0.5)	(0.4)	(0.5)	(0.2)	(0.2)	(0.4)
Y90~93	.025	.030	.023	.025	.033	.039	.013	.014
	(2.2)	(2.2)	(1.9)	(2.0)	(2.1)	(2.3)	(1.0)	(1.0)
Observations	17,572	17,572	34,653	34,653	27,126	27,126	25,099	25,099
Mean Dep. Var.	.338	.338	.368	.368	.410	.410	.301	.301
LR statistic	1214.0	2448.9	4637.0	8229.1	3318.4	6922.4	1926.9	3125.1
Significance	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

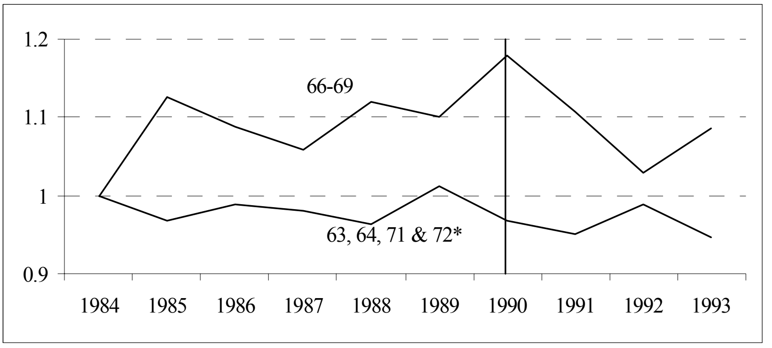
Marginal effects given for each model. The *t* statistics are given in parentheses. Models include five regional dummy variables not shown. See text and Table 2 for details.

TABLE 4
Tobit Analysis of Annual Hours Worked

Variables	Age 66 to 69 & No Control Group		Control Group Age 63, 64, 71 & 72		Control Group Age 63 & 64		Control Group Age 71 & 72	
	No wage	Wage	No wage	Wage	No wage	Wage	No wage	Wage
C	512.1	−3756.4	696.2	−4608.3	795.1	−5113.5	442.9	−3438.6
	(11.5)	(38.1)	(25.2)	(65.3)	(23.7)	(66.1)	(8.6)	(35.5)
AGE	−57.8	18.2	−70.2	−4.5	−71.7	17.9	−51.0	7.3
	(11.5)	(3.9)	(46.1)	(3.1)	(13.1)	(3.9)	(11.5)	(1.7)
ED8	−162.5	315.2	−202.7	408.9	−230.2	456.0	−147.2	274.5
	(9.4)	(17.6)	(15.7)	(31.3)	(14.7)	(30.4)	(11.0)	(18.3)
ED11	−126.5	115.4	−158.4	170.5	−173.7	175.6	−119.8	114.5
	(7.1)	(7.0)	(11.8)	(14.3)	(10.8)	(12.8)	(8.5)	(8.4)
ED13	105.4	−188.2	95.6	−294.9	112.0	−311.6	84.9	−190.0
	(6.0)	(11.3)	(7.3)	(24.6)	(7.2)	(22.7)	(6.2)	(13.8)
ED16	295.1	−518.3	280.1	−787.4	333.2	−866.5	235.1	−483.9
	(14.9)	(21.6)	(19.1)	(44.0)	(19.0)	(43.1)	(15.1)	(23.3)
ED17	385.2	−769.2	407.3	−1116.5	447.8	−1205.8	347.6	−729.8
	(18.4)	(25.7)	(26.3)	(49.5)	(24.3)	(49.1)	(21.0)	(26.7)
MARRIED	56.2	−231.7	60.3	−348.1	84.4	−351.9	31.6	−246.7
	(2.6)	(11.6)	(3.9)	(24.7)	(4.6)	(22.1)	(1.8)	(14.6)
METRO	−55.9	−366.0	−39.8	−439.0	−37.5	−496.2	−52.2	−316.7
	(4.7)	(29.4)	(4.5)	(48.4)	(3.5)	(47.6)	(5.6)	(30.6)
LN~FY	6.4	7.7	8.1	8.8	7.8	8.1	7.4	9.0
	(2.6)	(3.6)	(4.7)	(6.1)	(3.9)	(5.0)	(3.8)	(5.1)
LN~PY	−45.5	−9.0	−51.3	−10.1	−61.2	−9.8	−37.1	−9.7
	(24.0)	(5.0)	(36.7)	(8.1)	(36.7)	(6.8)	(24.8)	(6.7)
STATEUER	−26.4	−24.3	−28.9	−24.8	−30.6	−25.3	−25.3	−23.9
	(7.6)	(7.9)	(11.1)	(11.3)	(9.9)	(9.9)	(9.2)	(9.5)
WAGE	—	1934.2	—	2497.4	—	2737.1	—	1764.7
		(47.1)		(79.5)		(81.5)		(45.8)
YEAR	−10.5	3.9	−13.5	−1.2	−17.4	0.1	−7.6	1.6
	(2.7)	(1.1)	(4.7)	(0.5)	(5.0)	(0.0)	(2.5)	(0.6)
A66~69	—	—	−65.9	4.2	−96.4	−71.0	−13.5	95.0
			(6.1)	(0.5)	(3.8)	(3.5)	(0.6)	(4.6)
Y90~93	—	—	15.2	39.5	22.0	10.5	4.0	12.9
*A66~69			(0.9)	(0.8)	(1.0)	(0.6)	(0.2)	(0.7)
Y90~93	38.0	44.8	36.5	34.8	45.7	46.6	22.2	20.2
	(1.7)	(2.3)	(1.9)	(2.2)	(1.9)	(2.4)	(1.0)	(1.0)
Observations	17,572	17,572	34,653	34,653	27,126	27,126	25,099	25,099
Mean hours	452.3	452.3	545.4	545.4	624.5	624.5	394.7	394.7
% who worked	.338	.338	.368	.368	.410	.410	.301	.301
Adjusted R^2	078	.271	.164	.395	.156	.424	.078	.229

Marginal effects given for each model. The *t* statistics are given in parenthesis. Models include five regional dummy variables not shown. See text and Table 2 for details.

FIGURE 3
Normalized Employment Rates for Select Age Groups by Year (in sample)



* Weighted average based on 61 percent 62 to 63 year olds & 39 percent 71 and 72 year olds.

*Y90~93*A66~69* are not found to be statistically significant at the .10 level. However, their positive values (.002 to .007) are consistent with theoretical expectations and the results shown in Table 2. The total increase in employment rates of 66- to 69-year-olds after 1990 is 3.0 percentage points (.025 + .005) in the model that uses 63, 64-, 71- and 72-year-olds as a control group. As a result, the hypothesis that an unobserved factor (changes in health, wealth, pensions or other changes) is the major reason for the increase in employment among individuals aged 66 to 69 is not rejected. That is to say, the change in the employment rate for 66-to 69-year-olds partially reflects an increase in the employment rate of 63 to 72 year olds.

A higher hourly wage rate has a positive impact on the employment decision of elderly males. A one percent increase in an individual's expected wage rate tends to increase the probability of employment by 1.1 percentage points among 66- to 69-year-olds (1.8 and 1.0 percentage points for 63- to 64-year olds and 71- to 72-year olds, respectively). Unfortunately, the wage effect is not useful for a natural comparison with the *Y90~93* dummy variable effect. A reduction from a 50-percent earnings test penalty to a 33-percent penalty will increase the natural log of wages by .29 (ln(*WAGE**.67) − ln(*WAGE**.50)). This would suggest that employment rates increase by over 30 percentage points, not 3.0 percentage points. The difference can be explained by the fact that only marginal wages can potentially increase by this large amount, not average wages. It is unlikely that employment decisions are based entirely on the new potential marginal wage rate. As such, the coefficients for *WAGE* and *Y90~93* are not directly comparable.

Human capital, as measured by educational attainment, is generally found to be a significant, direct predictor of employment for 66- to 69-year olds and for the various control groups. Additional schooling generally leads to a reduced probability of employment, when controlling for expected wages. In addition, education exerts a strong offsetting indirect effect through its impact on the expected hourly wage rate. It is demonstrated in the Appendix that higher educational attainment levels increase an individual's expected wage rate. In turn, higher wages increase the probability of an individual working. In total, additional schooling increases the likelihood of employment as is evident in Probit models that exclude wages.

Living in a metropolitan area, *ceteris paribus*, decreases the likelihood of working by 21.1 percentage points for 66- to 69-year-olds (decreases of 31.9 and 17.4 percentage points are observed for 63- to 64- and 71- to 72-year-olds, respectively). This result is partially offset by the fact that such individuals have higher expected market wage rates. Living in a state with a high unemployment rate tends to decrease the probability that an individual worked during the calendar year, as expected. Being married has a negative impact on employment when holding wage rates equal. As predicted by neoclassical labor supply theory, non-wage income of the respondent has a statistically significant negative impact on employment. The income of other family members is found to be a significant positive predictor of employment.

The Tobit models find a positive aggregate impact in annual hours worked when looking solely at 66- to 69-year-olds. The difference-in-differences models show a positive but statistically insignificant increase in hours worked by individuals aged 66 to 69 relative to those in the control group after the 1990 rule change (note the positive but statistically insignificant coefficients for *Y90~93*A66~69*).

In total, all twelve models with control groups (both Probit and Tobit models) find the coefficients for the *Y90~93*A66~69* variables not to be statistically significant at the .10 level. This suggests that the impact of the 1990 change in the earnings test penalty is not the major cause of the change in employment behavior of elderly men. The results are consistent with the previous findings in the literature. Nevertheless, it is worth noting that all the coefficients for *Y90~93*A66~69* are positive, as expected.

CONCLUSIONS AND PUBLIC POLICY IMPLICATIONS

By analyzing the impact of the 1990 reduction in Social Security's earnings test penalty on employment rates, the paper adds significantly to the literature. Most of the previous literature on the earnings test fails to address the issue of its impact on the employment decision.

The 1990 change in Social Security's earnings test penalty is shown to have a positive, but statistically insignificant, impact on the employment rate of 66- to 69-year-old men relative to those in a valid control groups. The rule change is also shown to have little, if any, impact on annual hours worked.

As with most research, further empirical work needs to be done. The analysis could be expanded to research the impact of the 1990 law change on the labor supply behavior of women and minority men. Research should be conducted to estimate how the removal of the earnings test penalty in 2000 for individuals age 65 to 69 may impact the employment decisions of the elderly. Further, an understanding of how the legislated increase in the "normal retirement age" to age 66 in 2008 (and to age 67 in 2020) will impact labor supply behavior of the elderly needs to be developed.

APPENDIX
Delayed Retirement Credit

Not all individuals choose to retire at the age at which they are first eligible for full retirement benefits, the "normal retirement age" or NRA (currently age 65). Individuals who retire prior to the NRA receive reduced Social Security benefits. Workers who retire after the NRA receive an increase in the amount of their future benefits. By delaying retirement, the worker becomes eligible to receive a delayed retirement credit, DRC. In short, beneficiaries forgo current benefits in exchange for higher future benefits. The DRC is in the process of being slowly increased from 3.0 percent for individuals turning 65 in the mid- and late 1980s to 8.0 percent in 2008. A DRC of 7 to 8 percent is meant to be actuarially neutral based on average life expectancies. Reimers and Honig [1993; 1996] and Friedberg [2000] find no effect on labor supply when the DRC was raised. Attempts were made in this research to find the labor supply impact of changes in the DRC but none was found, perhaps because changes in the DRC and in the earnings test penalty occurred simultaneously. Consequently, it is assumed that individuals do not respond to changes in future benefits through the DRC.[8]

WAGE EQUATION

Sample selection in the wage equation is limited to those individuals who actually worked at some point during the year because otherwise no wage is observed. Approximately two-thirds of individuals in the sample did not work. It is possible that those who worked have unmeasured traits that lead them to earn a higher-than-expected market wage rate. To account for this, Heckman's [1979] two-stage estimation procedure is used. The first stage involves constructing a Probit equation estimating the probability of an individual working and having an observed wage. This stage uses a limited array of variables including a dummy variable that indicates whether an individual collected Social Security benefits, a strong negative predictor of employment. The log-wage models exclude those with a suspect wage or suspect hours of labor supplied. This includes those with a wage below half the national minimum wage ($1.50 in 1983 dollars), those with a real wage rate above $100.00 (the top one percentile), and those reported to be working more than 3,120 hours in a year. The second stage regressions include all the independent variables listed Table A1, including Heckman's lambda term (the inverse Mills ratio). Wages are estimated for the entire sample using the results found in Table A1 (stage two), but dropping the lambda term. The procedure is repeated to estimate wages for each of the various age groups.

TABLE A1
Estimated Log Wage Equation Using Heckman's Two Step Process

Variables	Age 66 to 69 & No Control Group: No wage	Wage	Control Group Age 63, 64, 71 & 72: No wage	Wage	Control Group Age 63 & 64: No wage	Wage	Control Group Age 71 & 72: No wage	Wage
C	0.33	2.61	0.43	2.27	0.52	2.24	0.23	2.65
	(24.9)	(8.3)	(47.0)	(66.8)	(47.7)	(58.9)	(22.5)	(20.3)
LN~PY	−0.01	—	−0.02	—	−0.02	—	−0.01	—
	(12.7)		(17.9)		(17.8)		(12.9)	
SS~YN	−0.42	—	−0.53	—	−0.60	—	−0.35	—
	(36.8)		(71.4)		(67.5)		(38.5)	
AGE	—	−0.06	—	0.00	—	0.03	—	−0.08
		(0.5)		(0.1)		(1.8)		(2.0)
AGE~SQ	—	0.00	—	0.00	—	−0.01	—	0.00
		(0.2)		(0.7)		(2.3)		(1.7)
ED8	—	−0.23	—	−0.22	—	−0.22	—	−0.22
		(6.5)		(9.6)		(9.2)		(7.2)
ED11	—	−0.11	—	−0.11	—	−0.11	—	−0.12
		(3.1)		(4.9)		(4.4)		(3.7)
ED13	—	0.14	—	0.15	—	0.14	—	0.15
		(4.2)		(6.9)		(6.3)		(5.1)
ED16	—	0.37	—	0.38	—	0.38	—	0.37
		(10.5)		(16.8)		(16.0)		(11.7)
ED17	—	0.52	—	0.55	—	0.53	—	0.55
		(14.2)		(23.3)		(21.4)		(17.0)
MARRIED	—	0.15	—	0.16	—	0.15	—	0.16
		(4.9)		(8.2)		(7.5)		(5.8)
METRO	—	0.15	—	0.14	—	0.15	—	0.14
		(6.5)		(9.9)		(9.8)		(6.9)
YEAR	—	0.08	—	0.00	—	0.00	—	−0.01
		(2.0)		(1.6)		(1.9)		(1.6)
LAMBDA	—	−0.47	—	−0.35	—	−0.37	—	−0.45
		(14.6)		(19.4)		(19.8)		(14.0)
Observations	17,572	5,118	34,653	11,057	27,126	9,708	25,099	6,467
Mean Dep. Var.	.338	2.130	.368	2.190	.410	2.210	.301	2.114
Adjusted R−sq.	—	.161	—	.177	—	.182	—	.153
F−statistic	—	62.3	—	149.8	—	136.07	—	73.82
LR statistic	1686.3	—	6389.1	—	5832.292	—	1790.527	—
Significance	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

SS~YN is a dummy variable indicating an individual collected Social Security. AGE~SQ equals AGE squared. Coefficients for AGE and AGE~SQ are jointly significant at the .01 level for both age groups. Coefficients in Model (1) are the marginal effects. The *t* statistics are given in parentheses. Model (2) includes five regional dummy variables, which are not shown. See text and Table 1 for details.

NOTES

I would like to thank Rexford Santerre, Andrew Sum, and two anonymous referees for their useful comments and suggestions.

1. Both Gruber and Orszag [2000] and Friedberg [2000] examine the impact of the 1990 rule change on labor supply of the elderly in other indirect ways (see the literature review section).
2. Net marginal wage rate refers to the wage rate net of taxes and current Social Security benefits.
3. Calculation based on CPS data. Similar results found in Friedberg [2000], Leonesio [1990], and Bondar [1993].
4. If no hours constraint exists, then the earnings test would merely limit hours and not impact the employment decision.
5. The earnings test penalty was eliminated for 70 and 71 year olds in 1983.
6. The focus of this paper is on employment rates, not the labor force participation rate. The two are closely related due to the low unemployment rates of the elderly.
7. The independent variables used to perform this test are identical to those used in the single difference model excluding the time shift dummy variable, Y90~93.
8. In contrast Rubb [2002] finds increases in the DRC to have a modest positive impact on the labor supply behavior of men during the mid-1990s.

REFERENCES

Blinder, A. S., Gordon, R. H., and Wise, D. E. Reconsidering the Work Disincentive Effect of Social Security. *National Tax Journal*, 1980, 431-42.

____________ **.** Rhetoric and Reality in Social Security Analysis- A Rejoinder. *National Tax Journal*, 1981, 473-78.

Bondar, J. Beneficiaries Affected by the Annual Earnings Test, 1989. *Social Security Bulletin*, 1993, 20-28.

Burkhauser, R. V., and Turner, J. Can Twenty-Five Million Americans Be Wrong? —A Response to Blinder, Gordon, And Wise. *National Tax Journal*, 1981, 467-72.

Burtless G., and Moffitt, R. A. The Effect of Social Security Benefits on the Labor Supply of the Aged, in *Retirement and Economic Behavior*, edited by H. Aaron and G. Burtless. Washington, D.C., 1984, 209-36.

____________**.** The Joint Choice of Retirement Age and Post-Retirement Hours of Work. *Journal of Labor Economics*, 1985, 209-236.

Friedberg, L. The Labor Supply Effects of the Social Security Earnings Test. *The Review of Economics and Statistics*, February 2000, 48-63.

Gruber, J. and Orszag, P. Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt? National Bureau of Economic Research Working Paper 7923, September 2000.

Gustman, A. L., and Steinmeier, T. L. Changing The Social Security Rules for Work After 65. *Industrial and Labor Relations Review*, May 1991, 733-45.

Herz, D. E. Work After Early Retirement: An Increasing Trend among Men. *Monthly Labor Review*, 1995, 13-21.

Honig, M. and Reimers, C. Is It Worth Eliminating the Earnings Test? *AEA Papers and Proceedings*, May 1989, 103-7.

Leonesio, M. V. The Effects of the Social Security Earnings Test on the Labor-Market Activity of Older Americans: A Review of the Evidence. *Social Security Bulletin*, 1990, 2-19.

____________**.** Social Security and Older Workers. *Social Security Bulletin*, 1993, 47-57.

Meyer, B. D. Natural and Quasi-Experiments in Economics. *Journal of Business and Economic Statistics*, April 1995, 151-61.

Packard, M. D. The Earnings Test and the Short-Run Work Response to its Elimination. *Social Security Bulletin*, 1990, 2-16.

Reimers, C. and Honig, M. The Perceived Budget Constraint under Social Security: Evidence from Reentry Behavior. *Journal of Labor Economics*, January 1993, 184-205.

____________**.** Responses to Social Security by Men and Women: Myopic and Far-Sighted Behavior. *Journal of Human Resources,* Spring 1996, 359-82.

Rubb, S. U.S. Social Security Rules in the 1990s: A Natural Experiment in Myopic and Farsighted Behaviour. *Applied Economics Letters.* August 2002, 637-40.